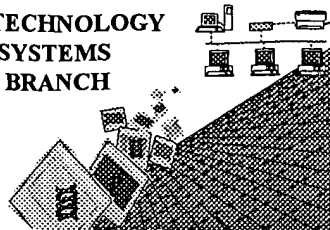BIOTECHNOLOGY
SYSTEMS
BRANCH

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information
Center (STIC) detected errors when processing the following computer readable
form:

Application Serial Number: _09/784,553A_

Source: _OIPE_

Date Processed by STIC: _11/7/2003_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE
   APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A
   NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
 PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
 PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
    http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is
a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the
United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm> , EFS Submission
   User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name,
   Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
          Or,
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two,
   2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office,
   Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 09/784,553 A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 _____ Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2 _____ Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 _____ Misaligned Amino Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4 _____ Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5 ✓ Variable Length

Sequence(s) __3__ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 _____ PatentIn 2.0 "bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7 _____ Skipped Sequences (OLD RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8 _____ Skipped Sequences (NEW RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9 _____ Use of n's or Xaa's (NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10 _____ Invalid <213> Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11 _____ Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12 _____ PatentIn 2.0 "bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 _____ Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001

OIPE

```
 3 <110> APPLICANT: Zhou, Ming-Ming
 4        Aggarwal, Aneel
 5        Verdin, Eric
 6        Ott, Melanie
 8 <120> TITLE OF INVENTION: Methods of Identifying Modulators of Bromodomains
10 <130> FILE REFERENCE: 2459-1-003CIP
12 <140> CURRENT APPLICATION NUMBER: 09/784,553A
13 <141> CURRENT FILING DATE: 2002-12-30
15 <150> PRIOR APPLICATION NUMBER: 09/510,314
16 <151> PRIOR FILING DATE: 2000-02-22
18 <160> NUMBER OF SEQ ID NOS: 59
20 <170> SOFTWARE: PatentIn version 3.0
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 3014
24 <212> TYPE: DNA
25 <213> ORGANISM: Homo sapiens
27 <400> SEQUENCE: 1
28 ggggccgcgt cgacgcggaa aagaggccgt ggggggcctc ccagcgctgg cagacaccgt    60
29 gaggctggca gccgccggca cgcacaccta gtccgcagtc ccgaggaaca tgtccgcagc   120
30 cagggcgcgg agcagagtcc cgggcaggag aaccaaggga gggcgtgtgc tgtggcggcg   180
31 gcggcagcgg cagcggagcc gctagtcccc tccctcctgg gggagcagct gccgccgctg   240
32 ccgccgccgc caccaccatc agcgcgcggg gcccggccag agcgagccgg gcgagcggcg   300
33 cgctaggggg agggcggggg cggggagggg ggtgggcgaa ggggcggga gggcgtgggg   360
34 ggagggtctc gctctcccga ctaccagagc ccgaggagа ccctggcggc ggcggcggcg   420
35 cctgacactc ggcgcctcct gccgtgctcc ggggcggcat gtccgaggct ggcggggccg   480
36 ggccgggcgg ctgcggggca ggagccgggg caggggccgg gcccggggcg ctgccccgc   540
37 agcctgcggc gcttccgccc gcgccccgc agggctcccc ctgcgccgct gccgccggg   600
38 gctcgggcgc ctgcggtccg gcgacggcag tggctgcagc gggcacggcc gaaggaccgg   660
39 gaggcggtgg ctcggcccga atcgccgtga agaaagcgca actacgctcc gctccgcggg   720
40 ccaagaaact ggagaaactc ggagtgtact ccgcctgcaa ggccgaggag tcttgtaaat   780
41 gtaatggctg gaaaaaccct aaccccctcac ccactccccc cagagccgac ctgcagcaaa   840
42 taattgtcag tctaacagaa tcctgtcgga gttgtagcca tgccctagct gctcatgttt   900
43 cccacctgga gaatgtgtca gaggaagaaa tgaacagact cctgggaata gtattggatg   960
44 tggaatatct ctttacctgt gtccacaagg aagaagatgc agataccaaa caagtttatt  1020
45 tctatctatt taagctcttg agaaagtcta ttttacaaag aggaaaacct gtggttgaag  1080
46 gctctttgga aaagaaaccc ccatttgaaa aacctagcat tgaacaggt gtgaataact  1140
47 ttgtgcagta caaatttagt cacctgccag caaaagaaag gcaaacaata gttgagttgg  1200
48 caaaaatgtt cctaaccgc atcaactatt ggcatctgga ggcaccatct caacgaagac  1260
49 tgcgatctcc caatgatgat atttctggat acaaagagaa ctacacaagg tggctgtgtt  1320
50 actgcaacgt gccacagttc tgcgacagtc tacctcggta cgaaaccaca caggtgtttg  1380
51 ggagaacatt gcttcgctcg gtcttcactg ttatgaggcg acaactcctg aacaagcaa  1440
52 gacaggaaaa agataaactg cctcttgaaa aacgaactct aatcctcact catttcccaa  1500
53 aatttctgtc catgctagaa gaagaagtat atagtcaaaa ctctcccatc tgggatcagg  1560
```

```
54 attttctctc agcctcttcc agaaccagcc agctaggcat ccaaacagtt atcaatccac   1620
55 ctcctgtggc tgggacaatt tcatacaatt caacctcatc ttcccttgag cagccaaacg   1680
56 cagggagcag cagtcctgcc tgcaaagcct cttctggact tgaggcaaac ccaggagaaa   1740.
57 agaggaaaat gactgattct catgttctgg aggaggccaa gaaaccccga gttatggggg   1800
58 atattccgat ggaattaatc aacgaggtta tgtctaccat cacggaccct gcagcaatgc   1860
59 ttggaccaga gaccaatttt ctgtcagcac actcggccag ggatgaggcg gcaaggttgg   1920
60 aagagcgcag gggtgtaatt gaatttcacg tggttggcaa ttccctcaac cagaaaccaa   1980
61 acaagaagat cctgatgtgg ctggttggcc tacagaacgt tttctcccac cagctgcccc   2040
62 gaatgccaaa agaatacatc acacggctcg tctttgaccc gaaacacaaa acccttgctt   2100
63 taattaaaga tggccgtgtt attggtggta tctgtttccg tatgttccca tctcaaggat   2160
64 tcacagagat tgtcttctgt gctgtaacct caaatgagca agtcaagggc tatggaacac   2220
65 acctgatgaa tcatttgaaa gaatatcaca taaagcatga catcctgaac ttcctcacat   2280
66 atgcagatga atatgcaatt ggatacttta agaaacaggg tttctccaaa gaaattaaaa   2340
67 tacctaaaac caaatatgtt ggctatatca aggattatga aggagccact ttaatgggat   2400
68 gtgagctaaa tccacggatc ccgtacacag aattttctgt catcattaaa aagcagaagg   2460
69 agataattaa aaaactgatt gaaagaaaac aggcacaaat tcgaaaagtt taccctggac   2520
70 tttcatgttt taaagatgga gttcgacaga ttcctataga aagcattcct ggaattagag   2580
71 agacaggctg gaaaccgagt ggaaaagaga aagtaaaga gcccagagac cctgaccagc   2640
72 tttacagcac gctcaagagc atcctccagc aggtgaagag ccatcaaagc gcttggccct   2700
73 tcatggaacc tgtgaagaga acagaagctc caggatatta tgaagttata aggttcccca   2760
74 tggatctgaa aaccatgagt gaacgcctca agaataggta ctacgtgtct aagaaattat   2820
75 tcatggcaga cttacagcga gtctttacca attgcaaaga gtacaacgcc gctgagagtg   2880
76 aatactacaa atgtgccaat atcctggaga aattcttctt cagtaaaatt aaggaagctg   2940
77 gattaattga caagtgattt tttttcccccc tctgcttctt agaaactcac caagcagtgt   3000
78 gcctaaagca aggt                                                     3014
80 <210> SEQ ID NO: 2
81 <211> LENGTH: 832
82 <212> TYPE: PRT
83 <213> ORGANISM: Homo sapiens
85 <400> SEQUENCE: 2
87 Met Ser Glu Ala Gly Gly Ala Gly Pro Gly Gly Cys Gly Ala Gly Ala
88 1               5                   10                  15
90 Gly Ala Gly Ala Gly Pro Gly Ala Leu Pro Pro Gln Pro Ala Ala Leu
91 .              20                  25                  30
93 Pro Pro Ala Pro Pro Gln Gly Ser Pro Cys Ala Ala Ala Ala Gly Gly
94         35                  40                  45
96 Ser Gly Ala Cys Gly Pro Ala Thr Ala Val Ala Ala Ala Gly Thr Ala
97      50                  55                  60
99 Glu Gly Pro Gly Gly Gly Gly Ser Ala Arg Ile Ala Val Lys Lys Ala
100 65                  70                  75                  80
102 Gln Leu Arg Ser Ala Pro Arg Ala Lys Lys Leu Glu Lys Leu Gly Val
103                 85                  90                  95
105 Tyr Ser Ala Cys Lys Ala Glu Glu Ser Cys Lys Cys Asn Gly Trp Lys
106              100                 105                 110
108 Asn Pro Asn Pro Ser Pro Thr Pro Pro Arg Ala Asp Leu Gln Gln Ile
109         115                 120                 125
111 Ile Val Ser Leu Thr Glu Ser Cys Arg Ser Cys Ser His Ala Leu Ala
112      130                 135                 140
114 Ala His Val Ser His Leu Glu Asn Val Ser Glu Glu Glu Met Asn Arg
```

```
115 145                    150                    155                    160
117 Leu Leu Gly Ile Val Leu Asp Val Glu Tyr Leu Phe Thr Cys Val His
118                    165                    170                    175
120 Lys Glu Glu Asp Ala Asp Thr Lys Gln Val Tyr Phe Tyr Leu Phe Lys
121                    180                    185                    190
123 Leu Leu Arg Lys Ser Ile Leu Gln Arg Gly Lys Pro Val Val Glu Gly
124             195                    200                    205
126 Ser Leu Glu Lys Lys Pro Pro Phe Glu Lys Pro Ser Ile Glu Gln Gly
127         210                    215                    220
129 Val Asn Asn Phe Val Gln Tyr Lys Phe Ser His Leu Pro Ala Lys Glu
130 225                    230                    235                    240
132 Arg Gln Thr Ile Val Glu Leu Ala Lys Met Phe Leu Asn Arg Ile Asn
133                    245                    250                    255
135 Tyr Trp His Leu Glu Ala Pro Ser Gln Arg Arg Leu Arg Ser Pro Asn
136                    260                    265                    270
138 Asp Asp Ile Ser Gly Tyr Lys Glu Asn Tyr Thr Arg Trp Leu Cys Tyr
139             275                    280                    285
141 Cys Asn Val Pro Gln Phe Cys Asp Ser Leu Pro Arg Tyr Glu Thr Thr
142         290                    295                    300
144 Gln Val Phe Gly Arg Thr Leu Leu Arg Ser Val Phe Thr Val Met Arg
145 305                    310                    315                    320
147 Arg Gln Leu Leu Glu Gln Ala Arg Gln Glu Lys Asp Lys Leu Pro Leu
148                    325                    330                    335
150 Glu Lys Arg Thr Leu Ile Leu Thr His Phe Pro Lys Phe Leu Ser Met
151                    340                    345                    350
153 Leu Glu Glu Glu Val Tyr Ser Gln Asn Ser Pro Ile Trp Asp Gln Asp
154             355                    360                    365
156 Phe Leu Ser Ala Ser Ser Arg Thr Ser Gln Leu Gly Ile Gln Thr Val
157         370                    375                    380
159 Ile Asn Pro Pro Pro Val Ala Gly Thr Ile Ser Tyr Asn Ser Thr Ser
160 385                    390                    395                    400
162 Ser Ser Leu Glu Gln Pro Asn Ala Gly Ser Ser Ser Pro Ala Cys Lys
163             405                    410                    415
165 Ala Ser Ser Gly Leu Glu Ala Asn Pro Gly Glu Lys Arg Lys Met Thr
166             420                    425                    430
168 Asp Ser His Val Leu Glu Glu Ala Lys Lys Pro Arg Val Met Gly Asp
169         435                    440                    445
171 Ile Pro Met Glu Leu Ile Asn Glu Val Met Ser Thr Ile Thr Asp Pro
172         450                    455                    460
174 Ala Ala Met Leu Gly Pro Glu Thr Asn Phe Leu Ser Ala His Ser Ala
175 465                    470                    475                    480
177 Arg Asp Glu Ala Ala Arg Leu Glu Glu Arg Arg Gly Val Ile Glu Phe
178             485                    490                    495
180 His Val Val Gly Asn Ser Leu Asn Gln Lys Pro Asn Lys Lys Ile Leu
181         500                    505                    510
183 Met Trp Leu Val Gly Leu Gln Asn Val Phe Ser His Gln Leu Pro Arg
184         515                    520                    525
186 Met Pro Lys Glu Tyr Ile Thr Arg Leu Val Phe Asp Pro Lys His Lys
187         530                    535                    540
```

RAW SEQUENCE LISTING
PATENT APPLICATION: **US/09/784,553A**

DATE: 01/17/2003
TIME: 11:03:34

Input Set : **A:\2459-1-003 Seqlist.txt**
Output Set: **N:\CRF4\01172003\I784553A.raw**

```
189 Thr Leu Ala Leu Ile Lys Asp Gly Arg Val Ile Gly Gly Ile Cys Phe
190 545             550             555         560
192 Arg Met Phe Pro Ser Gln Gly Phe Thr Glu Ile Val Phe Cys Ala Val
193             565             570             575
195 Thr Ser Asn Glu Gln Val Lys Gly Tyr Gly Thr His Leu Met Asn His
196             580             585             590
198 Leu Lys Glu Tyr His Ile Lys His Asp Ile Leu Asn Phe Leu Thr Tyr
199         595             600             605
201 Ala Asp Glu Tyr Ala Ile Gly Tyr Phe Lys Lys Gln Gly Phe Ser Lys
202     610             615             620
204 Glu Ile Lys Ile Pro Lys Thr Lys Tyr Val Gly Tyr Ile Lys Asp Tyr
205 625             630             635             640
207 Glu Gly Ala Thr Leu Met Gly Cys Glu Leu Asn Pro Arg Ile Pro Tyr
208             645             650             655
210 Thr Glu Phe Ser Val Ile Ile Lys Lys Gln Lys Glu Ile Ile Lys Lys
211             660             665             670
213 Leu Ile Glu Arg Lys Gln Ala Gln Ile Arg Lys Val Tyr Pro Gly Leu
214         675             680             685
216 Ser Cys Phe Lys Asp Gly Val Arg Gln Ile Pro Ile Glu Ser Ile Pro
217     690             695             700
219 Gly Ile Arg Glu Thr Gly Trp Lys Pro Ser Gly Lys Glu Lys Ser Lys
220 705             710             715             720
222 Glu Pro Arg Asp Pro Asp Gln Leu Tyr Ser Thr Leu Lys Ser Ile Leu
223             725             730             735
225 Gln Gln Val Lys Ser His Gln Ser Ala Trp Pro Phe Met Glu Pro Val
226             740             745             750
228 Lys Arg Thr Glu Ala Pro Gly Tyr Tyr Glu Val Ile Arg Phe Pro Met
229             755             760             765
231 Asp Leu Lys Thr Met Ser Glu Arg Leu Lys Asn Arg Tyr Tyr Val Ser
232     770             775             780
234 Lys Lys Leu Phe Met Ala Asp Leu Gln Arg Val Phe Thr Asn Cys Lys
235 785             790             795             800
237 Glu Tyr Asn Ala Ala Glu Ser Glu Tyr Tyr Lys Cys Ala Asn Ile Leu
238             805             810             815
240 Glu Lys Phe Phe Phe Ser Lys Ile Lys Glu Ala Gly Leu Ile Asp Lys
241             820             825             830
243 <210> SEQ ID NO: 3
244 <211> LENGTH: 12
245 <212> TYPE: PRT
246 <213> ORGANISM: artificial sequence
248 <220> FEATURE:
W--> 249 <221> NAME/KEY: X
250 <222> LOCATION: (2)..(2)
251 <223> OTHER INFORMATION: X is two to three amino acids. Each of these can be any
amino acid
253 <220> FEATURE:
W--> 254 <221> NAME/KEY: X
255 <222> LOCATION: (4)..(4)
256 <223> OTHER INFORMATION: The X is five to eight amino acids. Each of these can be any
amino acid
258 <220> FEATURE:
```

see p. 9 for error explanation
variable length not permitted
see item 5 on Error Summary Sheet
same error — see item 5

RAW SEQUENCE LISTING   DATE: 01/17/2003
PATENT APPLICATION: US/09/784,553A   TIME: 11:03:34

Input Set : A:\2459-1-003 Seqlist.txt
Output Set: N:\CRF4\01172003\I784553A.raw

W--> 259 <221> NAME/KEY: X   OK
     260 <222> LOCATION: (5)..(5)
     261 <223> OTHER INFORMATION: X is a single amino acid that is either Pro, Lys, or His.
     263 <220> FEATURE:
W--> 264 <221> NAME/KEY: X   _ OK
     265 <222> LOCATION: (6)..(6)
     266 <223> OTHER INFORMATION: This X is any single amino acid.
     268 <220> FEATURE:
W--> 269 <221> NAME/KEY: X   OK
     270 <222> LOCATION: (8)..(8)
     271 <223> OTHER INFORMATION: This X is a single amino acid that can be either Tyr, Phe,
or His
     273 <220> FEATURE: invalid-see item 5
W--> 274 <221> NAME/KEY: X
     275 <222> LOCATION: (9)..(9)
     276 <223> OTHER INFORMATION: X is 5 amino acids Each of these can be any amino acid.
     279 <220> FEATURE:
W--> 280 <221> NAME/KEY: X
     281 <222> LOCATION: (11)..(11)
     282 <223> OTHER INFORMATION: X is a single amino acid that can be Met, Ile, or Val.
     284 <400> SEQUENCE: 3
W--> 286 Phe Xaa Pro Xaa Xaa Xaa Tyr Xaa Xaa Pro Xaa Asp
     287 1   5   10
     289 <210> SEQ ID NO: 4
     290 <211> LENGTH: 12
     291 <212> TYPE: PRT
     292 <213> ORGANISM: artificial sequence   see p. 9 for error Explanation
     294 <220> FEATURE:
W--> 295 <221> NAME/KEY: Xaa
     296 <222> LOCATION: (6)..(6)   OK
     297 <223> OTHER INFORMATION: The X represents an acetyl-lysine
     300 <400> SEQUENCE: 4
W--> 302 Ile Ser Tyr Gly Arg Xaa Lys Arg Arg Gln Arg Arg
     303 1   5   10
     305 <210> SEQ ID NO: 5
     306 <211> LENGTH: 14
     307 <212> TYPE: PRT
     308 <213> ORGANISM: artificial sequence   p. 9
     310 <220> FEATURE:
W--> 311 <221> NAME/KEY: X
     312 <222> LOCATION: (8)..(8)   OK
     313 <223> OTHER INFORMATION: The X represents an acetyl lysine.
     316 <400> SEQUENCE: 5
W--> 318 Ala Arg Lys Ser Thr Gly Gly Xaa Ala Pro Arg Lys Gln Leu
     319 1   5   10
     321 <210> SEQ ID NO: 6
     322 <211> LENGTH: 14
     323 <212> TYPE: PRT
     324 <213> ORGANISM: artificial sequence
     326 <220> FEATURE:

The types of errors shown exist throughout
the Sequence Listing. Please check subsequent
sequences for similar errors.

The types of errors shown exist throughout
the Sequence Listing. Please check subsequent
sequences for similar errors.

<210> SEQ ID NO 44
<211> LENGTH: 20
<212> TYPE: PRT
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: :

see p. 9 for error explanation

<400> SEQUENCE: 44

```
Trp Pro Phe Met Glu Pro Val Lys Arg Thr Glu Ala Pro Gly Tyr Tyr
1               5                   10                  15
Glu Val Ile Arg
            20
```

```
<210> SEQ ID NO 45
<211> LENGTH: 101
<212> TYPE: PRT
<213> ORGANISM: Human immunodeficiency virus type 1     OK
<400> SEQUENCE: 45
      Met Glu Pro Val Asp Pro Arg Leu Glu Pro Trp Lys His Pro Gly Ser
      1               5                   10                  15
      Gln Pro Lys Thr Ala Ser Asn Asn Cys Tyr Cys Lys Arg Cys Cys Leu
                  20                  25                  30
      His Cys Gln Val Cys Phe Thr Lys Lys Gly Leu Gly Ile Ser Tyr Gly
                  35                  40                  45
      Arg Lys Lys Arg Arg Gln Arg Arg Arg Ala Pro Gln Asp Ser Lys Thr
              50                  55                  60
      His Gln Val Ser Leu Ser Lys Gln Pro Ala Ser Gln Pro Arg Gly Asp
      65                  70                  75                  80
      Pro Thr Gly Pro Lys Glu Ser Lys Lys Lys Val Glu Arg Glu Thr Glu
                      85                  90                  95
      Thr Asp Pro Glu Asp
                  100
```

```
<210> SEQ ID NO 46      see p. 9
<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: artificial sequence
<220> FEATURE:
<221> NAME/KEY: X                    invalid — see item 5
<222> LOCATION: (5)..(5)
<223> OTHER INFORMATION: X is one to three amino acids. Each amino acid can be
any amino
      acid
<400> SEQUENCE: 46
      Tyr Gly Arg Lys Xaa Arg Gln
      1               5
```

```
<210> SEQ ID NO 47
<211> LENGTH: 10
<212> TYPE: PRT
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: :          see p. 9
<400> SEQUENCE: 47
      Ser Tyr Gly Arg Lys Lys Arg Arg Gln Arg
      1               5                   10
```

```
<210> SEQ ID NO 48
<211> LENGTH: 10
<212> TYPE: PRT
<213> ORGANISM: artificial      see p. 9
<220> FEATURE:
<221> NAME/KEY: X
<222> LOCATION: (2)..(2)              invalid — see item 5
<223> OTHER INFORMATION: X is two to four amino acids.  Each of these can be any
amino aci
      d
<220> FEATURE:
<221> NAME/KEY: X
<222> LOCATION: (4)..(4)
```

<223> OTHER INFORMATION: X is two to four amino acids. Each of these can be any amino aci
    d

<220> FEATURE:
<221> NAME/KEY: X
<222> LOCATION: (6)..(6)
<223> OTHER INFORMATION: X is two to four amino acids. Each of these can be any amino aci
    d

<220> FEATURE:
<221> NAME/KEY: X
<222> LOCATION: (8)..(8)
<223> OTHER INFORMATION: X is one to three amino acids. Each of these can be any amino ac
    id

<220> FEATURE:
<221> NAME/KEY: X
<222> LOCATION: (10)..(10)
<223> OTHER INFORMATION: X is a single amino acid that is either Ile, Leu, Met, or Val.
<400> SEQUENCE: 48
    Phe Xaa Val Xaa Glu Xaa Tyr Xaa Val Xaa
    1   5   10

<210> SEQ ID NO 49
<211> LENGTH: 62
<212> TYPE: PRT
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: :
<400> SEQUENCE: 49
    Phe Met Glu Pro Val Lys Arg Thr Glu Ala Pro Gly Tyr Tyr Glu Val
    1   5   10   15

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 01/17/2003
PATENT APPLICATION: **US/09/784,553A**   TIME: 11:03:35

Input Set : **A:\2459-1-003 Seqlist.txt**
Output Set: **N:\CRF4\01172003\I784553A.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; Xaa Pos. 2,4,5,6,8,9,11
Seq#:4; Xaa Pos. 6
Seq#:5; Xaa Pos. 8
Seq#:6; Xaa Pos. 8
Seq#:43; Xaa Pos. 1,2,4,6,7,8,10,11,13
Seq#:46; Xaa Pos. 5
Seq#:48; Xaa Pos. 2,4,6,9,10
Seq#:50; Xaa Pos. 5
Seq#:51; Xaa Pos. 5
Seq#:52; Xaa Pos. 5
Seq#:53; Xaa Pos. 5
Seq#:54; Xaa Pos. 5
Seq#:55; Xaa Pos. 5
Seq#:56; Xaa Pos. 5
Seq#:57; Xaa Pos. 6
Seq#:58; Xaa Pos. 7
Seq#:59; Xaa Pos. 16

Invalid Line Length:
The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:34; Line(s) 1117
Seq#:35; Line(s) 1145

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:48

Use of <220> Feature(NEW RULES):
Sequence(s)__are missing the <220> Feature and associated headings.
Use of <220> to <223> is MANDATORY if <213> ORGANISM is "Artificial Sequence"
or"Unknown". Please explain source of genetic material in <220> to <223>
section (See "Federal Register," 6/01/98, Vol. 63, No. 104,pp.29631-32)
(Sec.1.823 of new Rules)

Seq#:44,47,49, 3,4 (maybe more)
46,48

Input Set : **A:\2459-1-003 Seqlist.txt**
Output Set: **N:\CRF4\01172003\I784553A.raw**

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:249 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:254 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:259 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:264 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:269 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:274 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:280 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:3
L:286 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0
L:295 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:4
L:302 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0
L:311 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:5
L:318 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0
L:327 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:6
L:334 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0
L:1387 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:43
L:1393 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:43
L:1400 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:43
L:1407 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:43
L:1413 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:43
L:1419 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:43
L:1425 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:43
L:1431 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:43
L:1438 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43 after pos.:0
L:1446 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ#:44, <213>
ORGANISM:artificial sequence
L:1446 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ#:44, <213>
ORGANISM:artificial sequence
L:1446 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:44,Line#:1446
L:1488 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:46
L:1496 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46 after pos.:0
L:1504 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ#:47, <213>
ORGANISM:artificial sequence
L:1504 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ#:47, <213>
ORGANISM:artificial sequence
L:1504 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:47,Line#:1504
L:1515 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:48
L:1522 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:48
L:1529 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:48
L:1536 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:48
L:1543 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:48
L:1550 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48 after pos.:0
L:1558 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ#:49, <213>
ORGANISM:artificial sequence
L:1558 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ#:49, <213>
ORGANISM:artificial sequence
L:1558 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:49,Line#:1558
L:1578 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:50
L:1585 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50 after pos.:0
L:1594 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:51
L:1601 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51 after pos.:0
L:1610 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:52
L:1617 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52 after pos.:0

L:1626 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:53

VERIFICATION SUMMARY　　　　　　　　　　DATE: 01/17/2003
PATENT APPLICATION:　**US/09/784,553A**　　　TIME: 11:03:35

Input Set : **A:\2459-1-003 Seqlist.txt**
Output Set: **N:\CRF4\01172003\I784553A.raw**

```
L:1633 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:53 after pos.:0
L:1642 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:54
L:1649 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:54 after pos.:0
L:1658 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:55
L:1665 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55 after pos.:0
L:1674 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:56
L:1681 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56 after pos.:0
L:1687 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:57
L:1690 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:57
L:1697 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57 after pos.:0
L:1706 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:58
L:1713 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58 after pos.:0
L:1722 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:59
L:1729 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:59 after pos.:0
```